Deutsches Institut für
Wirtschaftsforschung

DIW BERLIN

Discussion Papers

1103

Katharina Pijnenburg • Konstantin A. Kholodilin

Do Regions with Entrepreneurial Neighbors Perform Better?
A Spatial Econometric Approach for German Regions

Berlin, February 2011

IMPRESSUM

DIW Berlin
German Institute for Economic Research
Mohrenstr. 58
10117 Berlin

Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
http://www.diw.de

ISSN print edition 1433-0210
ISSN electronic edition 1619-4535

Papers can be downloaded free of charge from the DIW Berlin website:
http://www.diw.de/discussionpapers

Discussion Papers of DIW Berlin are indexed in RePEc and SSRN:
http://ideas.repec.org/s/diw/diwwpp.html
http://www.ssrn.com/link/DIW-Berlin-German-Inst-Econ-Res.html

Do Regions with Entrepreneurial Neighbors Perform Better?

A Spatial Econometric Approach for German Regions

Katharina Pijnenburg* Konstantin A. Kholodilin

DIW Berlin DIW Berlin

this version: February 7, 2010

Abstract

We use a neoclassical production function to analyze the effects of knowledge spillovers via entrepreneurship on economic performance of 337 German districts. To take the spatial dependence structure of the data into account, we estimate a spatial Durbin model. We highlight the importance of the choice of the appropriate weight matrix. We find positive knowledge spillover effects via entrepreneurship within a certain region. Between regions, entrepreneurship as a vehicle by which knowledge spills over and contributes to economic performance depends largely on the choice of the weight matrix. We see this as evidence for regionally bounded knowledge spillover effects via entrepreneurship.

Keywords: entrepreneurship capital, regional output, spatial weight matrix

JEL Classification: C21, M13, R11

*This paper represents the authors' personal opinions and does not necessarily reflect the views of DIW Berlin. Contact: German Institute for Economic Research, Mohrenstraße 58, 10117 Berlin, Germany, Phone: +49 30 89789-589, Fax: +49 30 89789-200, E-mail: kpijnenburg@diw.de

1 Introduction

Despite a vast literature about economic performance, a universally accepted model has not been found yet. The Mankiw et al. (1992) human capital augmented Solow model produced by now probably the most convincing empirical results. But an emerging literature suggests that economic development is highly related to the abundance of small entrepreneurial firms. In this literature start-ups constitute an important link between knowledge creation and knowledge commercialization, which will generate economic output. Acs and Armington (2004) found for the U.S. that firm start-ups are an important vehicle, by which knowledge spills over and contributes to economic growth. The Sutter (2010) results show that the commercial introduction of knowledge via firm start-ups has a larger effect on economic growth than pure knowledge creation. For West German regions Audretsch and Fritsch (2002) found that start-up rates have a positive impact on growth rates in the 1990s. Moreover, Audretsch and Keilbach (2004) discovered in their analysis of German regions that the start-up rate had a significant effect on economic output in 1992. More generally, Fischer and Nijkamp (2009) state that regional change is the result of entrepreneurial activity where innovations play a key role.

The knowledge spillover theory of entrepreneurship (Acs et al. (2009)), which underlies those studies focuses on individuals with endowments of new economic knowledge.[1] This knowledge was previously generated in an university or an incumbent firm, in which the individual is working. The expected value of this new idea can be higher for this individual than for the decision maker in the university or in the incumbent firm. If the expected return is sufficiently high and the costs of starting a new business sufficiently low, the individual will enter the market and start her own business. The so created start-up is the vehicle with which knowledge spills over from the source of knowledge production to a new firm that will commercialize it.[2]

However, the decision to become an entrepreneur depends as well on a positive entrepreneurial environment. This introduces the concept of entrepreneurship capital. Audretsch and Keilbach (2004) defined it “as a region's endowment with factors conducive to the creation of new business." Those factors are, for example, individuals that are willing to start a new business, an innovative milieu, networks, institutions, which facilitate bureaucratic steps involved and institutions like banks that are willing to share risks. In the theory of knowledge spillover of entrepreneurship, the knowledge stock is only a necessary condition for economic growth. The theory

[1]Economic knowledge is knowledge which holds commercial opportunity.

[2]This is in contrast to Romer (1990) and Lucas (1988), where knowledge exogenously spills over between firms.

gives a deeper understanding of the essential role of entrepreneurship for economic development.

In our analysis of regional economic growth in Germany we take this essential role of entrepreneurship into account. We estimate a neoclassical production function model augmented by entrepreneurship capital. We choose a regional context as it is widely recognized that the region is the fundamental basis of economic life (Capello and Nijkamp (2009)). Furthermore, empirical studies found that knowledge spillovers tend to be locally bounded (Acs et al. (1994), Audretsch and Feldman (1996)). Since we test the hypothesis that entrepreneurship serves as a conduit for knowledge spillovers, the regional level is appropriate.

The novelty of this paper is that we take the spatial dimension of the data explicitly into account. Even if it has widely been accepted that empirical analysis on a regional level needs a spatial econometric context to account for spatial dependence in the data, for entrepreneurship capital and economic output this has to our knowledge only been done in Sutter (2010) for the U.S. We estimate a spatial Durbin model where we put special emphasis on the creation and the choice of the weight matrix. This is another aspect that has not received much attention even in the spatial econometric literature. For the case of Germany we could not find a single paper where the choice of the weight matrix played a role. The weight matrix determines to what extent region i affects region j and vice versa. In spatial econometrics inference and estimates depend on the weight matrix used. Different weight matrix specifications may have an important impact on coefficient estimates (LeSage and Fischer (2008)). The spatial econometric estimation method will allow us not only to find out how large is the knowledge spillover effect via entrepreneurship on economic output in a certain region, but also how large is the spillover effect coming from neighboring regions. This will allow us to answer the question posed in the title. Namely, if a region performs better if it has entrepreneurial neighbors. We find evidence in favor of knowledge spillover effects via entrepreneurship within a certain region. The evidence for knowledge spillovers via entrepreneurship between regions is weak. Having entrepreneurial neighbors does not significantly affect a regions performance. Human capital is found to exert a positive effect on economic output but human capital of the neighboring regions has a large negative impact on economic output. This delivers evidence for the presence of a brain drain.

The analysis clearly confirms the presence of a spatial dependence structure. A failure to account for it would result in biased estimates. We further show that the results vary depending on the choice of the weight matrix.

The remainder of the paper is organized as follows. In section 2, we start with the theoretical model and the data description. Section 3 will treat spatial estimation issues. In detail we explain the spatial Durbin model, the Bayesian estimation method, the correct coefficient interpretation and the creation and comparison of the weight matrices. Section 4 presents our empirical results, section 5 concludes.

2 Model

To analyze the knowledge spillover effects via entrepreneurship, we consider a neo-classical production function, which not only includes the standard variables physical capital (K) and human capital (H) as explanatory variables but also entrepreneurship capital (E): equation (1). In this way we follow the approach of Audretsch and Keilbach (2004).

$$Y = F(K, H, E, L) \tag{1}$$

The variables are divided by labor (L) so we work with per effective unit of labor variables and have thereby productivity expressions, y, k, h, e. With a Cobb-Douglas specification of the production function we get:

$$y_i = a k_i^{\alpha_K} h_i^{\alpha_H} e_i^{\alpha_E}, \tag{2}$$

where $i = 1, ..., n$ denotes regions and a represents the state of the technology. Taking logs yields:

$$\ln y_i = \ln a + \alpha_K \ln k_i + \alpha_H \ln h_i + \alpha_E \ln e_i. \tag{3}$$

We estimate this theoretical model with data for 337 German NUTS 3 regions (Nomenclature of Territorial Units for Statistics) for the year 2004. We choose this year as the data we used for entrepreneurship capital are only available for a time span of four years, namely 1997-2000, 2001-2004, and 2005-2008 and the data we used for economic output, for example, are only available up until 2007. The most recent overlap is therefore given in 2004. All variables are per working age population ratios.

The dependent variable economic output, y, was measured by gross value added at basic prices, physical capital, k, was calculated with the perpetual inventory method. This procedure allows to compute the stock of physical capital (K) as the weighted sum of past investments (I) in manufacturing and mining ($K_t = I_t + (1 - \delta)K_{t-1}$). For the calculation we chose that data on investment in 1995 as initial capital stock. We assumed a depreciation rate, δ, of five percent (Barro and Sala-i-Martin (1995), Chew and Tan (1999)). We know that this captures only part of the total investments and could result in misleading coefficients, but data on gross fixed capital formation is not published on a regional level. Further, we used the share of employees with technical college or university degree in the working age population to measure human capital, h. This definition is in line with what is quite often used in the literature (Barro and Lee (1993), Fischer et al. (2009), LeSage and Fischer (2008)). Following Audretsch and Keilbach (2004), entrepreneurship capital, e, is approximated by start-up rates in knowledge intense areas. We use this variable as entrepreneurship capital is a variable that cannot be observed but should manifest itself in high start-up rates. The Mannheimer Unternehmenspanel provides those start-up rates for several categories. We defined six categories as being knowledge intense and aggregated them (Table 1). We only used knowledge intense start-ups as we are interested in knowledge spillovers via entrepreneurship, which are most likely to occur in those areas (Acs et al. (2009)).

Variable	**Description**	**Year**	**Source**
Output	Gross value added at basic prices	2004	Eurostat
Physical capital	Investment in manufacturing and mining, Perpetual inventory method, Delta = 5 %	1995-2004	German Statistical Office (destatis)
Human capital	Share of employees with technical college or university degree on working age population	2004	German Statistical Office (destatis)
Entrepreneurship capital	Start up rates in the following areas: - Cutting-edge technology manufacturing, - High-technology manufacturing, - Technology-intense services - Skill-intense services - ICT software supply and consultancy - ICT trade and renting	2001-2004	Mannheimer Unternehmenspanel (ZEW - Center for European Economic Research)

Table 1: Data description

3 Spatial Econometric Modeling

3.1 Spatial Durbin Model

As LeSage and Pace (2010) noted, data collected from regions are often not independent. This spatial dependence requires a special estimation method because neglecting this structure would result in biased estimates. There are four well known spatial econometric specifications, namely the spatial lag model (SAR), which includes a spatial lag of the dependent variable, the spatial error model (SEM), which includes a spatial lag in the error term, the spatial autoregressive moving average model (SARMA), which includes a spatial lag of the dependent variable and in the error term, and the spatial Durbin model (SDM), which includes a spatial lag of the dependent and the explanatory variables.

The spatial lag $\sum_{j=1}^{n} W_{\mathrm{ij}} y_j$ is the weighted average of the spatially lagged variables of the neighboring regions. W is the spatial weight matrix of dimension $n \times n$. If two regions i and j are spatially related, the element $w_{ij} \neq 0$, otherwise $w_{ij} = 0$. By convention a region cannot be a neighbor to itself, $w_{ii} = 0$. To simplify interpretation the weight matrix is usually row-standardized, so that the row sums are equal to one. The spatial lag operator then corresponds to the weighted average of neighboring observations.

LeSage and Pace (2009) point out that the SDM is the only model that will produce unbiased estimates no matter which of the mentioned data generating processes is underlying. This is why we choose the spatial Durbin model as appropriate estimation specification (Equation (4)). This model further nests the spatial lag and the spatial error model, i.e. models involving dependence in the error term and in the dependent variable.

$$\begin{aligned} y &= a\iota_n + \rho W y + \mathrm{X}\beta \; + W\mathrm{X}\gamma + \epsilon \qquad (4) \\ \epsilon &\sim N\left(0, \sigma^2 I_n\right) \end{aligned}$$

In this equation, ρ measures the strength of the spatial lag dependence of the dependent variable, Wy. γ measures the strength of the spatial lag dependence of the explanatory variable Wx. This spatial model specification applied to our neoclassical production function model yields the following equation which we are going to estimate.

$$\begin{aligned} y &= a\iota_n + \rho\, W y + \beta_1 k + \beta_2 h + \beta_3 e + W k \gamma_1 + W h \gamma_2 + W e \gamma_3 + \epsilon \qquad (5) \\ \epsilon &\sim N\left(0, \sigma^2 I_n\right) \end{aligned}$$

Where y is the dependent variable economic output, Wy is the spatial lag of economic output, Wk is the spatial lag of the independent variable physical capital, Wh is the spatial lag of the independent variable human capital, and We is the spatial lag of the independent variable entrepreneurship capital. This specification allows us not only to explicitly account for spatial dependence in the data but we further will get insight about regional spillovers of the three explanatory variables.

3.2 Estimation Method

Spatial models can be estimated with maximum likelihood or Bayesian estimation methods. Our focus is on the comparison of different weight matrices. Tests, like the likelihood ratio test, that would use the log likelihood function values to compare the models can only be applied for nested models. Two models with different weight matrices can not be considered as being nested. That is why we use Bayesian estimation. The Bayesian posterior model probabilities allow model comparison even for non-nested models.

Bayesian estimation in general is centered around posterior probabilities. $P(\theta|D)$ is the so called posterior probability of the parameters, θ, given the data, D, and reflects the belief about the parameters after collecting the data.

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)} \tag{6}$$

The posterior distribution represents an update of the prior distribution given the data. $P(D|\theta)$ is the model likelihood and $P(\theta)$ is the prior distribution of the parameters and reflect previous knowledge or uncertainty prior to observing the data. The probability of the data $P(D)$ is not of great interest as it does not involve the parameters θ. Bayesian inference about parameters is entirely based on the posterior distribution $P(\theta|D)$.

We apply the Bayesian Markov Chain Monte Carlo approach to estimate the parameters α, β_r, γ_r, ρ, and σ^2, where r is 1 to 3, and stands for the explanatory variables. By applying this method, we work with a large random sample from the posterior distribution and not with the precise analytical form of the density. A large sample of the posterior probability distribution allows us to approximate the analytical form of the probability density. In a first step we assign prior distributions to the parameters of our spatial Durbin model. We follow LeSage and Pace (2009) and assign the normal prior to α, β, and γ, the inverse gamma prior to σ^2, and the uniform prior to ρ.

The Bayesian Markov Chain Monte Carlo works as follows: In a first step we begin with arbitrary parameter values $\beta\,(0)$, $\gamma\,(0)$, $\sigma^2(0)$, and $\rho(0)$ and sample sequentially from the conditional distributions. At first we sample for β using the normal distribution and taking the arbitrary parameter values for γ, ρ, and σ^2. The sampled parameter vector is $\beta(1)$ and replaces $\beta(0)$. Next, we sample for γ using the normal distribution and $\beta\,(1)$, $\sigma^2(0)$, and $\rho\,(0)$. Afterwards we sample for σ^2 using the inverse gamma distribution and $\beta\,(1)$, $\gamma\,(1)$, and $\rho\,(0)$. Finally, we sample for ρ.

Those four steps are repeated 7500 times. We assume that the sampler achieves its steady state after 2500 draws. That is why the first 2500 draws are excluded.

The last 5000 draws are interpreted as coming from the posterior distribution. We use the large sample of parameter draws from the posterior distribution to make inference about α, β_r, γ_r, ρ, and σ^2. Inference is based on statistics like the mean and the standard deviation of the parameter sample.
We also account for heteroscedasticity in the data by extending the above described Markov Chain Monte Carlo estimation by variance scalars that can accommodate non-constant variance of the error term:

$$\epsilon \sim N\left(0, \sigma^2 V\right), \tag{7}$$

where V is a diagonal matrix containing the parameters $(\upsilon_1, \upsilon_2, ..., \upsilon_n)$, which are unknown and need to be estimated. We assign a chi-squared prior distribution, $\chi^2(s)/s$, to the υ_i terms. The Markov Chain Monte Carlo sampling scheme is extended by an additional conditional distribution for the variance scalars. We need to account for these new parameters in the model and adjust the conditional posterior distributions for the other parameters. Following LeSage and Pace (2009), s of the chi-squared prior distribution is set to 4, as this is consistent with a prior belief in non-constant variance and outliers.

3.3 Coefficient Interpretation

The coefficients of Bayesian estimation can not be interpreted as marginal effects. This comes from the spatial dependence structure in the data. A change in the explanatory variable of region i will affect the region i itself, which is called a direct impact, and potentially this change will also affect all other regions, which is called an indirect impact. Spatial econometric models are able to capture this effects. LeSage and Pace (2009) explain how to calculate those summary marginal measures of impact. To understand the effects of an expanded information set coming from neighboring regions, we look at the data generating process of the spatial Durbin model.

$$\begin{aligned} (I_n - \rho \mathrm{W})\, y &= \alpha \iota_n + \mathrm{X}\beta + \mathrm{WX}\gamma + \epsilon \\ y &= (I_n - \rho W)^{-1} \iota_n \alpha + (I_n - \rho W)^{-1} X\beta + \\ &\quad (I_n - \rho W)^{-1} WX\gamma + (I_n - \rho W)^{-1} \epsilon, \end{aligned} \tag{8}$$

where

$$(I_n - \rho W)^{-1} = I_n + \rho W + \rho^2 W^2 + \rho^3 W^3 + ... \tag{9}$$

Part of the right-hand side can be reformulated as

$$\begin{aligned}(I_n - \rho W)^{-1} X\beta + (I_n - \rho W)^{-1} WX\gamma &= \sum_{r=1}^{3} (I_n - \rho W)^{-1} (I_n \beta_r + W\gamma_r)\, x_r \\ &= \sum_{r=1}^{3} S_r(W) x_r, \end{aligned} \tag{10}$$

again, r stands for the explanatory variables. Using this expression, the data generating process can be written as:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sum_{r=1}^{3} \begin{pmatrix} S_r(W)_{11} & S_r(W)_{12} & \cdots & S_r(W)_{1n} \\ S_r(W)_{21} & S_r(W)_{22} & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ S_r(W)_{\mathrm{n1}} & S_r(W)_{\mathrm{n2}} & \cdots & S_r(W)_{\mathrm{nn}} \end{pmatrix} \begin{pmatrix} x_{1r} \\ x_{2r} \\ \vdots \\ x_{\mathrm{nr}} \end{pmatrix} + (I_n - \rho W)^{-1} (\iota_n \alpha + \epsilon). \tag{11}$$

We see that the derivative of y_i with respect to x_{jr} does not equal β_r but:

$$\frac{\partial y_i}{\partial x_{\mathrm{jr}}} = S_r(W)_{\mathrm{ij}}. \tag{12}$$

This means that a change in the explanatory variable for region j may have an effect on the dependent variable of all other regions. Furthermore, the derivative of y_i with respect to x_{ir} does not equal β_r but:

$$\frac{\partial y_i}{\partial x_{\mathrm{ir}}} = S_r(W)_{\mathrm{ii}}. \tag{13}$$

This implies that there are feedback loops, as region i affects region j and region j, in turn, affects region i.
The matrix $S_r(W)$ can be used to calculate the above mentioned direct and indirect impacts. The average direct effect is calculated as the average of the diagonal of the matrix $S_r(W)$, which can be calculated using the trace of the matrix:

$$\bar{M}(r)_{\mathrm{direct}} = \frac{1}{n} \mathrm{tr}(S_r(W)). \tag{14}$$

The average total impact is calculated as the average of all derivatives of y_i with respect to x_{jr} for any i and j:

$$\bar{M}(r)_{\text{total}} = \frac{1}{n}\iota_n'(S_r(W))\iota_n. \tag{15}$$

The average indirect effect is calculated as the average total effect minus the average direct effect

$$\bar{M}(r)_{\text{indirect}} = \bar{M}(r)_{\text{total}} - \bar{M}(r)_{\text{direct}}. \tag{16}$$

In such a way we are able to calculate the direct, the indirect, and the total average impact effects of the variables. The average direct effect is the one that comes from the same region i. The average indirect effect is the one that comes from the other regions $j \neq i$. The average total effect is the sum of the direct and the indirect effect.

During the Markov Chain Monte Carlo sampling we can construct those three summary measures. We simply use at each pass through the Markov Chain Monte Carlo sampling loop the sampled parameters α, β, γ, and ρ to calculate the multiplier term $S_r(W) = (I_n - \rho W)^{-1}(I_n\beta_r + W\gamma_r)$. In this way we obtain with each draw the direct effect, the total effect and by subtracting the direct effect from the total effect, the indirect effect. Then we can construct the entire posterior distribution for the three types of marginal effects using the 5000 saved draws.

3.4 Spatial Weight Matrix Comparison

We choose the spatial Durbin model but there is still model uncertainty due to diverse specification possibilities of the weight matrix. Finding the weight matrix that best reflects the spatial dependence is a key element of spatial econometric analysis. There are several methods to create a weight matrix. The weights could be based on geographical, technological, economic, demographical or political distance. The weight matrix based on geographical proximity can be constructed using either of two approaches: a binary measure of contiguity and a continuous measure of distance. Contiguity measures take the word neighbor in its proper sense. If two regions i and j share a common border they are considered to be first order contiguous and the value of 1 is assigned to w_{ij}. Higher orders of contiguity can be considered as well. Contiguity of order c assigns the value of 1 to regions which share a common border with a region that is a $(c-1)$ order contiguous region.

Continuous distance measures assign a value $w_{ij} > 0$ to regions that are in a certain distance to region i. This value is obtained using, for example, geographical coordi-

nates of two regions i and j together with a distance decay function. Distance decay functions have the effect of reducing the influence of regions j on region i as the distance between them increases. When using distance measures most studies rely on geographical proximity. LeSage and Fischer (2008) use geodesic distance, road travel time distances for cars and drive time distances for heavy goods vehicles.
Parent and LeSage (2008) do not only rely on geographical distance to create their weight matrix, but also on technological distance. They measure the distance between technological fields by using patent activity occurring between regions in the same field of technology defined by the International Patent Classification. LeSage and Polasek (2008) incorporate prior information about commodity flows transported by road and rail into the spatial connectivity structure and Beck et al. (2006) use volume of trade flows.
Case et al. (1993) construct weight matrices that are based on economic and demographical proximity, where they use per-capita income for economic proximity and the percentage of the population that is black for demographical proximity. Bhattacharjee and Jensen-Butler (2006) propose to use the data in the model to estimate the spatial weight matrix that is consistent with the observed pattern of spatial dependence.
For our analysis we created 43 row standardized spatial weight matrices which can be arranged by four types of geographical distance measures. For the first seven weight matrices we used the c nearest neighbors, where $c = 3, ..., 9$. We started with c equal to 3 as German Kreise are quite small and it appears unrealistic if one creates a weight matrix, where only the neighboring region ($c = 1$) or the regions that are neighbors and neighbors to the neighbors ($c = 2$) are included. If a district is a c nearest neighbor, the value one is assigned to that region. This procedure results in a binary contiguity matrix.
The direct distance, calculated between the centers of two regions together with the cut-off distance b was used for the next twelve weight matrices. The cut-off distance $b = 50$ determines that only regions that are within a 50 km radius around the region under examination have an impact. We choose b to be 50, 100, 150, 200, 250, and 300 and created a weight matrix for each distance twice: once with the power distance decay function ($w_{\mathrm{ij}} = \frac{1}{d_{\mathrm{ij}}{}^{p}}$, where we set p equal to one, and d_{ij} is the distance between region i and region j) and once with the exponential distance decay function ($w_{\mathrm{ij}} = \frac{1}{\exp^{\mathrm{pd_{ij}}}}$, again, we set p equal to one).
The next twelve matrices where created similar to the previous twelve, but that time by using the road distance between regions instead of the direct distance. We choose the most fuel-efficient route, not the shortest or the fastest one. Even if the infrastructure is quite well developed in Germany, trucks need to cover often a longer distance on the road than the direct distance. For instance, the direct distance be-

tween Flensburg and Lübeck, which are two cities located in the north of Germany, was calculated with 132 km. The road distance which was taken from the route planning site viamichelin.de is 160 km. If we take one city in the north of Germany and another in the south the difference between direct and road distance becomes more pronounced in absolute terms. The direct distance between Flensburg and Lindau (Bodensee) was calculated with 800 km, whereas the road distance is 948 km.

The last twelve weight matrices account for the time on the road that one needs to cover the distance between two regions. As cut-off time d, we used $d = 1, ..., 6$ hours. The duration may give us additional insight as it varies even if we have quite similar distances. The road distance from Flensburg to Göttingen is 419, that from Flensburg to the Grafschaft Bentheim 412. For the first distance the duration is 5 hours and 30 minutes for the second only 4 hours and 15 minutes.

Bayesian model comparison allows to find out which of the 43 weight matrices fits the data best. Therefore, we look at the posterior model probabilities. But first we need to specify prior probabilities for each model. We assign to each model the same probability, namely $\frac{1}{m}$, where m is the number of different models. Together with the prior distributions for the parameters we can calculate posterior model probabilities. The joint posterior for the models M and the parameters θ is given by:

$$P(M, \theta|D) = \frac{\pi(M)\pi(\theta|M)P(D|\theta, M)}{P(D)}. \tag{17}$$

The marginal posterior probability of the models, where we integrated over θ is:

$$P(M|D) = \int P(M, \theta|y) d\theta. \tag{18}$$

Those posterior probabilities are then directly compared. The model with the highest posterior probability fits the sample data best.

4 Estimation Results

4.1 Coefficient Comparison

Table 2 shows that the direct, the indirect, and the total marginal effect estimates of the spatial Durbin models differ depending on the choice of the weight matrix. We present here the minimum and the maximum value of those coefficients that were significant in the estimation at least at a ten percent level. The coefficients of the direct marginal effects do not vary much across models. This can also be seen in the figures 1, 2, and 3 in the appendix, where we plotted the direct, the indirect, and the total effects of physical capital (k), human capital (h), and entrepreneurship capital (e) of all estimations. The dot stands for the total marginal effect. Again, only coefficients, which were significant, are presented.

The size of the indirect effect varies quite strongly for the three variables over the different estimations. This can be seen from the smallest and the largest coefficients in Table 2 and from figures 1 through 3. For physical capital we found the largest indirects effects in those models where the weight matrices were created with the direct or road distance with a cut-off point of 50, 100, and 150 km and in those models where we used the duration distance with a cut-off time at one hour. For the duration time we observe the same for entrepreneurship capital, furthermore large coefficients can be found with the direct distance and a cut-off value of 50 and 100 km, and for the road distance with a cut-off distance at 50 km. It appears that regional spillovers of physical and entrepreneurship capital are especially pronounced if only close regions have a weight unequal to zero in the weight matrix.

For human capital there is no clear pattern, large negative marginal effects can be found for different distances. We could draw the conclusion that regional human capital spillovers are not especially pronounced for close regions, but are also present if more distant regions are included. It appears that regional spillovers are more regionally bounded for physical and entrepreneurship capital than for human capital.

The results further suggest that finding a positive knowledge spillover effect via entrepreneurship across regions depends to a large extent on the choice of the weight matrix. Only 20 out of 43 coefficients of the indirect effects of entrepreneurship capital are significant.

We further used the boxplot to visually summarize the coefficients of the different estimations (Figures 4 to 6). The top of the box represents the 75th and the bottom the 25th percentile. The line in the box is the median of the coefficients. The smallest and the largest coefficients are connected with the box by whyskers. Outliers are represented by a dot. We see that for the direct effects the coefficients for each of the three variables lie close together. For the indirect effect there is obviously larger variation, especially for human and entrepreneurship capital, and some coefficients are considered as outliers. In the estimation those are not significant. The same is true for the total effects.

4.2 Model Comparison

These different coefficient results, emphasize the importance of a careful choice of the weight matrix. We used diverse measures of physical distance to create the

	Direct		
	Min	Max	significant coefficients
physical capital k	0.1061	0.1547	43
human capital h	0.2444	0.3242	43
entrepreneurship capital e	0.0711	0.1898	43
	Indirect		
	Min	Max	significant coefficients
physical capital k	0.0756	0.4071	27
human capital h	-1.736	-0.165	32
entrepreneurship capital e	0.0714	0.3932	20
	Total		
	Min	Max	significant coefficients
physical capital k	0.2302	0.514	29
human capital h	-1.317	0.0959	28
entrepreneurship capital e	0.2589	0.6257	29

Table 2: Coefficient comparison

weight matrices. As described above, posterior model probabilities are then used to compare the different model specifications. Those can be found in Table 3, where the 6 highest probabilities are reported. The Bayesian model probabilities point with 0.47 to the road power distance with a cut-off distance at 100 km. This weight matrix therefore best fits the sample data. The direct power distance with a cut-off distance at 50 km has the second highest probability with 0.37. Therefore, the

Distance	Model Probability
Road distance, cut-off 100 km, power distance function	0.472
Direct distance, cut-off 50 km, power distance function	0.365
Direct distance, cut-off 100 km, power distance function	0.148
Road distance, cut-off 150 km, power distance function	0.006
Duration distance, cut-off 2 hours, power distance function	0.006
Duration distance, cut-off 2 hours, exponential distance function	0.002

Table 3: Posterior model probabilities

spatial Durbin model estimation, on which we should rely, is the one that uses the weight matrix based on road power distance with a cut-off distance at 100 km. The results further suggest that in the case of Germany spatial effects are strongest, when only quite close neighbors are considered.

4.3 Results of the Final Model

The results of the final model can be found in Table 4. We see that the Bayesian

W = Road distance, cut-off distance 100 km, power distance function Bayesian Spatial Durbin Model				
		Direct	**Indirect**	**Total**
constant	-2.36 ***			
physical capital k (β_1)	0.12 ***	0.13 ***	0.31 **	0.43 ***
human capital h (β_2)	0.32 ***	0.30 ***	-0.67 ***	-0.38 ***
entrepreneurship capital e (β_3)	0.12 ***	0.13 ***	0.21	0.34 **
W-k (γ_1)	0.05			
W-h (γ_2)	-0.46 ***			
W-e (γ_3)	0.00			
ρ	0.62 ***			

*** Statistically significant at one percent level
** Statistically significant at five percent level

Table 4: Estimation output

coefficient estimates are quite similar to the direct summary effects. Moreover, we see that we cannot interpret the coefficients of spatially lagged explanatory variables (W-k, W-h, W-e) as spillovers. The true spillovers are represented by the indirect effects and are quite different. As the indirect effect is significant for physical capital and human capital, and the total effect for all three variables, it implies the necessity of accounting for spatial effects. Furthermore, ρ, which describes the strength of the spatial dependence of the dependent variable is large and highly significant. If we would neglect the spatial dependence structure, which is clearly present, we would get biased estimates.
For physical capital all three impact measures are positive and significant. As the direct impact is smaller than the indirect impact, we conclude that the neighbors stock of physical capital has a larger impact on the performance of region i than its own physical capital stock. LeSage and Fischer (2008) found a small negative indirect effect of physical capital but do not give further explanations for this result. As all variables are in logs, the total effect of 0.43 implies that a 10 percent increase in all regions physical capital stock will increase output by 4.3 percent. If we would not consider the effect of the neighboring regions in our estimation, we would largely underestimate the marginal effect of physical capital.
The largest direct effect is found for human capital. A ten percent increase in human capital of region i will increase economic output of region i by three percent. Fischer et al. (2009) and LeSage and Fischer (2008) found a smaller direct effect of

human capital for European regions in the magnitude of 0.12 in their analysis of regional labor productivity and regional growth, respectively. To the contrary, the indirect marginal effect of human capital is negative. That means that a ten percent increase in human capital of the regions that according to the weight matrix have an influence on region i will reduce economic output of region i by almost seven percent. Compared to the LeSage and Fischer (2008) and Fischer et al. (2009) studies this effect is quite large. The authors found an indirect effect of -0.11 and -0.20, respectively. The difference in the amplitude of the indirect effect is probably due to the use of NUTS 2 regions in their studies and NUTS 3 regions in ours. This implies that the indirect effect of human capital is more pronounced if smaller regions are analyzed. Furthermore, it is possible that human capital spillovers between regions are larger in a national context than in an international one. An explanation why human capital of the other regions affects the output of region i negatively is quite obvious. The increase in human capital in region j could imply that working conditions are better there than in region i. Better conditions attract workers from region i and result in an output reduction of that region (Nistor (2009)). Our results can be interpreted as evidence for the brain drain. Due to this large negative indirect effect, the total effect for human capital is negative as well. The result is not in line with the studies of LeSage and Fischer (2008) and Fischer et al. (2009), who found insignificant total effects of human capital. Our results suggest that regions should have a large interest in improving working conditions as there is high competition over human capital.

The coefficient of the direct marginal effect of entrepreneurship capital (0.13) points to the presence of knowledge spillovers via entrepreneurship. A ten percent increase in entrepreneurship capital will result in more knowledge commercialization and an increase in output by 1.3 percent. This is in line with the Audretsch and Keilbach (2004) results. They found a coefficient of entrepreneurship capital in a neoclassical production function model of 0.12. But they do not account for spatial dependence in their analysis of West German regions. To our knowledge, the only study, which uses entrepreneurship in a spatial Durbin model is the one by Sutter (2010). He found a direct effect of entrepreneurship on factor productivity of 0.5 for U.S. states. Regarding the indirect impact of entrepreneurship capital we could not find a significant effect. But the total effect, which is the sum of the direct effect and the indirect effect, is significant. Therefore, we cannot ignore the positive indirect effect of entrepreneurship capital on economic output. Compared to the direct effect, the indirect effect (0.21) is quite large. But it is smaller than the indirect effect of entrepreneurship (0.66) found by Sutter (2010) for the U.S. The indirect effect of entrepreneurship capital can be interpreted as positive knowledge spillover via entrepreneurship from neighboring regions. If a region increases its endowment with entrepreneurship capital, neighboring regions will benefit from it. This may come from entrepreneurs who use, for example, the network and the financial assistance in one region but create their business in another region. Thereby knowledge spills over not only from an existing firm into the economy but also between regions. But the effect of knowledge spillover via entrepreneurship between regions is much less important. That can be concluded from the fact that only 20 out of 43 indirect marginal effects calculated with the different weight matrices (Figure 3) are significant. If we come back to our initial question we have to answer it with a no. A region will not necessarily perform better if it has entrepreneurial neighbors.

5 Conclusion

In our analysis of regional economic performance of German regions, we estimated a neoclassical production function with physical capital, human capital, and entrepreneurship capital as explanatory variables. We were especially interested in the role of firm start-ups as link between knowledge creation and knowledge commercialization. We used the spatial Durbin model as econometric model because it allows to take into account the spatial dependence structure of the data. We put special emphasis on the creation and comparison of diverse weight matrices. A weight matrix determines to what extent region j affects region i and vice versa. The weight matrix, which was created with the road distance within a distance of 100 km together with the power distance decay function, was found to be the matrix, which best mirrors the true spatial dependence structure. We found this result by comparing posterior model probabilities, which we calculated in a Bayesian Markov Chain Monte Carlo estimation.

We found significant knowledge spillovers via entrepreneurship within a region. There is only weak evidence for the presence of knowledge spillovers via entrepreneurship between regions. We conclude that knowledge spillovers via entrepreneurship are locally bounded and that it does not help much for a regions performance to have entrepreneurial neighbors. We further found a positive direct effect of human capital on economic output. The effect of human capital of the neighboring regions j on the economic performance of region i was found to be negative. This is interpreted as evidence for the brain drain. This indirect effect is found to be even larger than the direct positive effect.

This analysis provides evidence for the importance of entrepreneurship for economic performance. The pure creation of knowledge is a necessary condition for higher economic output but in order to commercialize this knowledge entrepreneurship is needed. Regarding the econometric setting we have shown the importance of the use of a spatial econometric model and the necessity of a careful choice of the weight matrix.

Appendix

Coefficient Comparison

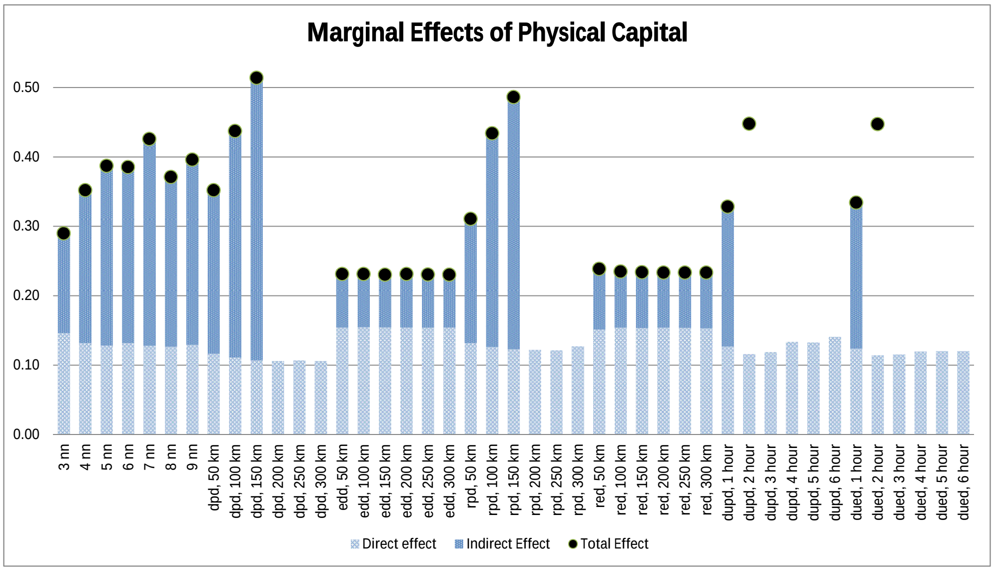

Figure 1: Coefficient comparison for different weight matrices: physical capital

Notes: Coefficients are only plotted if significant, nn = nearest neighbor, dpd = direct power distance, edd = exponential direct distance, rpd = road power distance, red = road exponential distance, dupd = duration power distance, dued = duration exponential distance

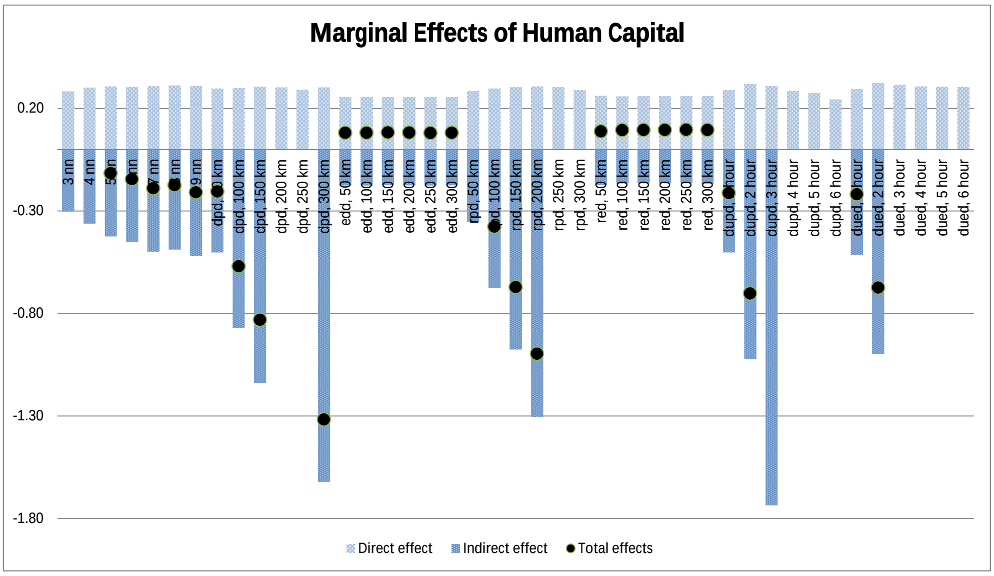

Figure 2: Coefficient comparison for different weight matrices: human capital

Notes: Coefficients are only plotted if significant, nn = nearest neighbor, dpd = direct power distance, edd = exponential direct distance, rpd = road power distance, red = road exponential distance, dupd = duration power distance, dued = duration exponential distance

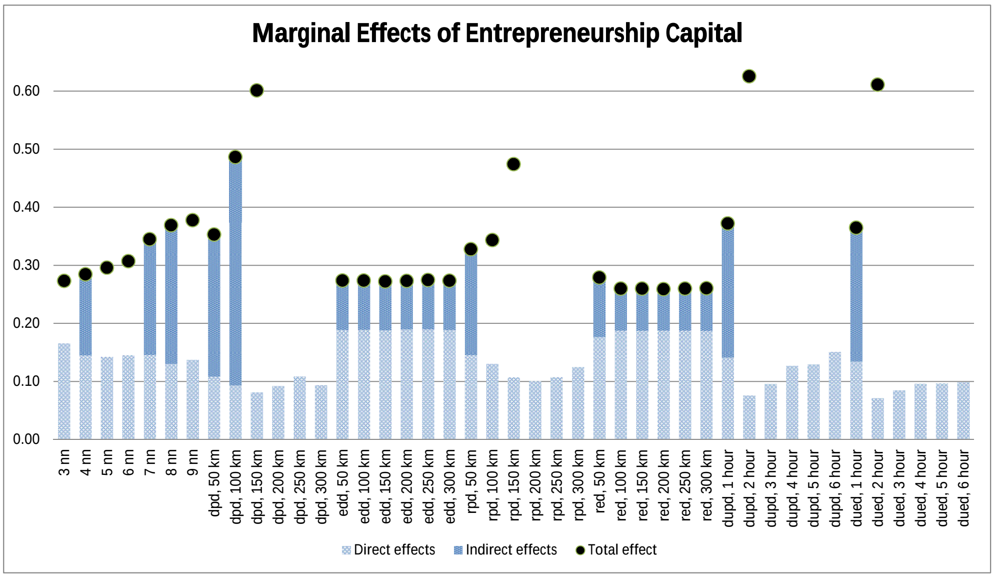

Figure 3: Coefficient comparison for different weight matrices: entrepreneurship capital

Notes: Coefficients are only plotted if significant, nn = nearest neighbor, dpd = direct power distance, edd = exponential direct distance, rpd = road power distance, red = road exponential distance, dupd = duration power distance, dued = duration exponential distance

Boxplots

Direct Marginal Effects

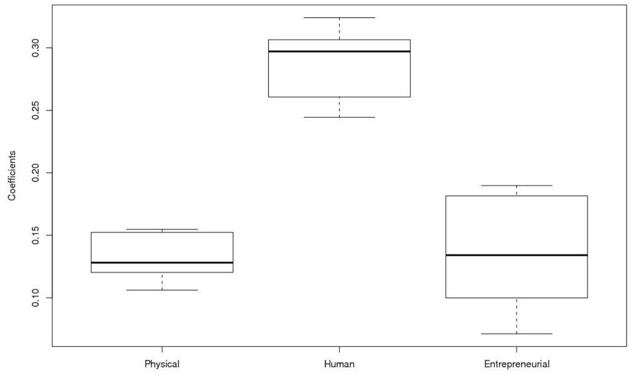

Figure 4: Boxplot for physical capital, human capital, and entrepreneurship capital for the direct effects

Indirect Marginal Effects

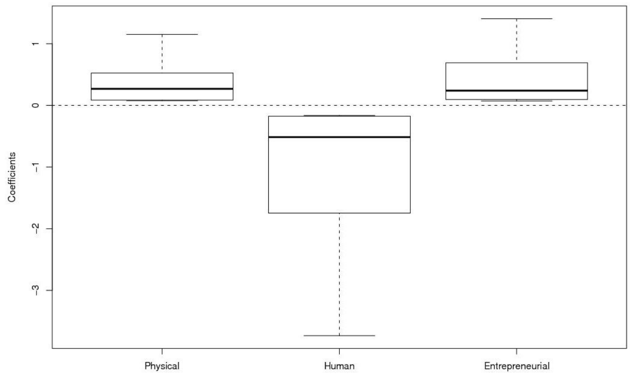

Figure 5: Boxplot for physical capital, human capital, and entrepreneurship capital for the indirect effects

Total Marginal Effects

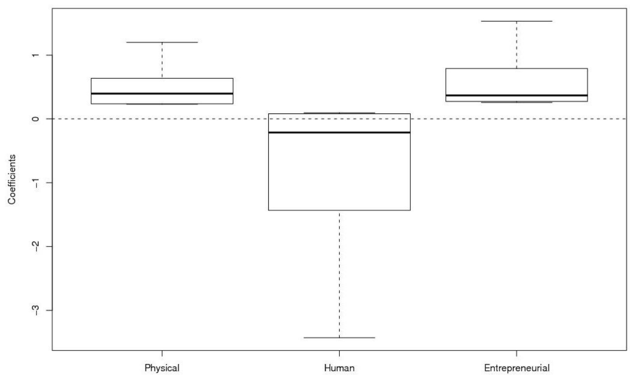

Figure 6: Boxplot for physical capital, human capital, and entrepreneurship capital for the total effects

References

Acs, Z., Armington, C., 2004. Employment growth and entrepreneurial activity in cities. Regional Studies 38 (8), 911–927.

Acs, Z., Audretsch, D., Feldman, M., 1994. R&D spillovers and innovative activity. Managerial and Decision Economics 15 (2), 131–138.

Acs, Z., Braunerhjelm, P., Audretsch, D., Carlsson, B., 2009. The knowledge spillover theory of entrepreneurship. Small Business Economics 32 (1), 15–30.

Audretsch, D., Feldman, M., 1996. R&D Spillovers and the Geography of Innovation and Production. The American Economic Review 86 (3), 630–640.

Audretsch, D., Fritsch, M., 2002. Growth regimes over time and space. Regional Studies 36 (2), 113–124.

Audretsch, D., Keilbach, M., 2004. Entrepreneurship capital and economic performance. Regional Studies 38 (8), 949–959.

Barro, R., Lee, J., 1993. International comparisons of educational attainment. Journal of Monetary Economics 32 (3), 363–394.

Barro, R., Sala-i-Martin, X., 1995. Economic Growth. McGraw-Hill.

Beck, N., Gleditsch, K., Beardsley, K., 2006. Space is more than geography: Using spatial econometrics in the study of political economy. International Studies Quarterly 50 (1), 27–44.

Bhattacharjee, A., Jensen-Butler, C., 2006. Estimation of spatial weights matrix, with an application to diffusion in housing demand. Centre for Research into Industry, Enterprise, Finance, and the Firm Discussion Paper 519.

Capello, R., Nijkamp, P., 2009. Introduction: Regional growth and development theories in the twenty-first century - Recent theoretical advances and future challenges in Handbook of Regional Growth and Development Theories. Edward Elgar Publishing Limited.

Case, A. C., Rosen, H. S., Hines, J. R., 1993. Budget spillovers and fiscal policy interdependence: Evidence from the states. Journal of public economics 52 (3), 285–307.

Chew, R., Tan, L., 1999. Human capital formation as an engine of growth: The East Asian experience. Institute of Southeast Asian Studies.

Fischer, M., Bartkowska, M., Riedl, A., Sardadvar, S., Kunnert, A., 2009. The impact of human capital on regional labor productivity in Europe. Letters in Spatial and Resource Sciences 2 (2), 97–108.

Fischer, M., Nijkamp, P., 2009. Entrepreneurship and Regional Development in Handbook of Regional Growth and Development Theories. Edward Elgar Publishing Limited.

LeSage, J., Fischer, M., 2008. Spatial growth regressions: Model specification, estimation and interpretation. Spatial Economic Analysis 3 (3), 275–304.

LeSage, J., Pace, R., 2009. Introduction to spatial econometrics. Chapman & Hall/CRC.

LeSage, J., Pace, R., 2010. Spatial econometric models. Handbook of Applied Spatial Analysis, 355–376.

LeSage, J., Polasek, W., 2008. Incorporating transportation network structure in spatial econometric models of commodity flows. Spatial Economic Analysis 3 (2), 225–245.

Lucas, R., 1988. On the mechanics of economic development. Journal of Monetary Economics 22 (1), 3–42.

Mankiw, N., Romer, D., Weil, D., 1992. A contribution to the empirics of economic growth. The Quarterly Journal of Economics 107 (2), 407–437.

Nistor, A., 2009. Assessing the Effectiveness of Human Capital Investments on the Regional Unemployment Rate in the United States: 1990 and 2000. International Regional Science Review 32 (1), 65.

Parent, O., LeSage, J., 2008. Using the variance structure of the conditional autoregressive spatial specification to model knowledge spillovers. Journal of Applied Econometrics 23 (2), 235–256.

Romer, P., 1990. Endogenous technological change. Journal of Political Economy 98 (S5), S71–S102.

Sutter, R., 2010. The Psychology of Entrepreneurship and the Technological Frontier–A Spatial Econometric Analysis of Regional Entrepreneurship in the United States. Ph.D. thesis, George Mason University.